دانشگاه یزد
دانشکده کشاورزی و منابع طبیعی
گروه محیط زیست

پایان‌نامه
برای دریافت درجه کارشناسی ارشد
مهندسی منابع طبیعی-محیط زیست

مقایسه‌ی کاربرد روش‌های شبکه‌ی عصبی مصنوعی و ماشین بردار پشتیبان برای پیش‌بینی آلودگی هوای شهری: (مطالعه‌ی موردی شهر تهران)

استاد راهنما: دکتر فرهاد نژادکورکی
استاد مشاور: دکتر سیدمحمد تقی المدرسی

پژوهش و نگارش: ادریس پایا

اسفند ۱۳۹۳

کلیه‌ی حقوق مادی و معنوی مترتب بر نتایج مطالعات، ابتکارات و نوآوری‌های ناشی از تحقیق موضوع این پایان‌نامه/ رساله متعلق به دانشگاه یزد است و هرگونه استفاده از نتایج علمی و عملی از این پایان نامه / رساله برای تولید دانش فنی، ثبت اختراع، ثبت اثر بدیع هنری، همچنین چاپ و تکثیر، نسخه برداری، ترجمه و اقتباس و ارائه مقاله در سمینارها و مجلات علمی از این پایان‌نامه / رساله منوط به موافقت کتبی دانشگاه یزد است.

با سپاس از سه وجود مقدس:

آنان که ناتوان شدند تا ما به توانایی برسیم...

موهایشان سپید شد تا ما روسفید شویم...

و عاشقانه سوختند تا گرمابخش وجود ما و روشنگر راهمان باشند...

پدر انمان

مادر انمان

استاد انمان

از اســتاد محترم، آقای **دکتر فرهاد نژادکورکی** که در کمال ســعه‌صــدر، با حســن خلق و فروتنی، از هیچ کمکی در انجام این تحقیق دریغ ننمودند و زحمت راهنمایی این رساله را بر عهده داشــتند و همچنین از اســتاد محترم، آقای **دکتر ســید محمد تقی المدرســی** که زحمت مشاوره این رساله را بر عهده داشتند، کمال تشکر و قدردانی دارم.

چکیده

آلودگی هوا یکی از مسائل مهم زیست محیطی در شهر تهران محسوب می‌شود. و برای بسیاری از ساکنین شهرهای بزرگ مانند تهرانیم موضوع نگران کننده می‌باشد. در سال‌های اخیر با افزایش چشمگیر وسائل نقلیه میزان آلودگی در شهر تهران افزایش قابل تجهی اشته است. به همین جهت، بررسی وضعیت آلودگی هوای این شهر از بعد آلاینده‌های مختلف می‌تواند در برنامه ریزی‌های جامعی که در راستای کنترل و کاهش آلودگی هوا در این شهر انجام می‌پذیرد، بسیار مؤثر باشد. پیش‌بینی آلاینده‌های هوا این امکان را بوجود می‌آورد که با آگاهی از قبل در مورد وضعیت آلاینده‌های هوا اقدامات لازم جهت مقابله با آن را محیا کرد. در همین راستا روش‌های زیادی برای پیش‌بینی غلظت آلاینده‌ها وجود دارد که روش‌های هوش مصنوعی می‌باشد. در این پژوهش از دو مدل شبکه‌عصبی مصنوعی شامل پرسپترون چندلایه، توابع شعاع مبنا و مدل رگرسیون ماشین بردار پشتیبان برای پیش‌بینی غلظت در سه ساعت پیش‌رو استفاده شد. پیش‌بینی با استفاده چهار نوع متفاوت ورودی صورت گرفت. ابتدا پیش‌بینی بر اساس ارزش‌های گذشته هر غلظت در پنج ایستگاه در سطح شهر تهران انجام گرفت. سپس با استفاده از کل داده‌های موجود شامل پارامترهای هواشناسی و آلودگی هوا به عنوان ورودی مدل‌های پیش‌بینی کننده استفاده شد. پس از آن با استفاده از پرامترهای هواشناسی به دو صورت کل پارامترهای موجود و با کاستن دو پارامتر از کل پارامترهای موجود پیش‌بینی غلظت آلاینده‌ها انجام گرفت. از توابع خودهمبستگی جزئی و همبستگی عرضی برای تعیین گام‌های زمانی موثر در مقدار خروجی استفاده شد. مقادیر حداقل و حداکثر R^2 به‌دست آمده در مجموع پنج ایستگاه بدون ورودی‌های خارجی برای غلظت‌های SO2، NO، NO2، COو O3 به ترتیب، 0.5، 0.3، و 0.73، 0.48و 0.53، 0.23و ۰,۷۱، 0.35و 0.86، 0.54، به دست آمد. این مقادیر زمانی که ورودی کل پارامترهای موجود استفاده شد برای SO2، NO، NO2، COو O3 به ترتیب ، 0.55، 0.28و 0.68، 0.54و 0.58، 0.28و 0.76، 0.38و 0.86، 0.53، به دست آمد. برای کل پارامترهای هواشناسی به‌عنوان ورودی برای غلظت‌های SO2، NO،

NO2، COو O3 به ترتیب 0.32، 0.11و 0.44، 0.16، 0.5، 0.14و 0.56، 0.14و 0.13 0.58به دست آمد. همچنین زمانی‌که تعداد پارامترهای هواشناسی به اندازه دو پارامتر کاسته شد مقادیر حداقل و حداکثر R^2 بدست آمده با استفاده از سه مدل پیش‌بینی کننده برای غلظت‌های SO2، NO، NO2، COو O3 به ترتیب 0.31، 0.01، 0.52، 0.05، 0.34و 0.07و 0.59و 0.07و 0.64و 0.13به می‌باشد. با توجه به نتایج به دست آمده هر سه مدل در پیش‌بینی غلظت آلاینده‌ها از عملکردهای تقریباً یکسانی برخوردار بودند. با انجام آنالیز واریانس با ضریب اطمینان ۹۵٪ مشخص شد که اختلاف معنی‌داری در صحت پیش‌بینی مدل‌ها با هر چهار نوع ورودی وجود نداشت. همچنین با کاهش تعداد پارامترهای هواشناسی به عنوان ورودی مدل، صحت پیش‌بینی مدل رگرسیون ماشین بردار پشتیبان نسبت به مدل‌های شبکه عصبی افزایش نیافته است.

کلمات کلیدی: آلودگی، هوش مصنوعی، همبستگی جزئی، همبستگی، تحلیل واریانس

فهرست مطالب

فصل اول: ۱

مقدمه و کلیات ۱

۱-۱ مقدمه ۲

۱-۲ بیان مسئله ۳

۱-۳ اهداف ۵

۱-۴ فرضیات ۵

۱-۵ کلیات ۶

۱-۵-۱ آلودگی هوا ۶

۱-۵-۲ منابع آلودگی هوا ۶

۱-۵-۲-۱ آلاینده‌های مهم و اثرات آن‌ها ۷

۱-۵-۲-۱-۱ هیدروکربن‌ها ۷

۱-۵-۲-۱-۲ مونواکسیدکربن ..... **Error! Bookmark not defined.**

۱-۵-۲-۱-۳ اکسیدهای گوگرد ۸

۱-۵-۲-۱-۴ اکسیدهای نیتروژن ۹

۱-۵-۲-۱-۵ اکسیدکننده‌های فوتوشیمیایی ۱۱

۱-۵-۲-۱-۶ باران‌های اسیدی ۱۱

۱-۵-۲-۲ اثرات عمومی آلودگی هوا ۱۲

۱-۵-۲-۲-۱ مستقیم ۱۲

۱-۵-۲-۲-۲ غیر مستقیم ..... **Error! Bookmark not defined.**

۱-۵-۲-۳ مدل‌سازی غلظت آلاینده‌های هوا ۱۳

۱-۵-۳ هوش مصنوعی ۱۴

۱-۵-۳-۱ شبکه عصبی مصنوعی ۱۵
۱-۵-۳-۲ ماشین بردار پشتیبان ۱۶
۲ فصل دوم: مرور منابع ۱۸
۲-۱ مقدمه ۱۹
۲-۲ بررسی مطالعات پیشین در رابطه با شبکه‌های عصبی مصنوعی ۱۹
۲-۳ بررسی مطالعات پیشین در رابطه با ماشین بردار پشتیبان ۲۸
۲-۴ بررسی مطالعات پیشین در مقایسه دو روش ۲۹
۲-۵ جمع‌بندی ۳۳
۳ فصل سوم: مواد و روش‌ها ۳۵
۳-۱ مقدمه ۳۵
۳-۲ موقعیت جغرافیایی منطقه مورد مطالعه ۳۵
۳-۳ شرایط آب و هوایی استان تهران ۳۶
۳-۴ آماده سازی داده‌ها ۳۶
۳-۴-۱ داده‌های مورد استفاده ۳۶
۳-۴-۲ تعیین تأخیر زمانی پارامترهای ورودی ۳۸
۳-۴-۲-۱ تابع خودهمبستگی جزئی برای تعیین تأخیر زمانی ۳۸
۳-۴-۲-۲ تابع همبستگی عرضی برای تعیین تأخیر زمانی ۳۸
۳-۵ معرفی مدل‌های مورد استفاده ۳۹
3-5-1 شبکه عصبی مصنوعی ۳۹
۳-۵-۲ ویژگی‌های شبکه عصبی مصنوعی ۴۰
۳-۵-۳ مدل نورون ۴۲
۳-۵-۳-۱ مدل تک ورودی ۴۲
۳-۵-۳-۲ توابع تحریک ۴۴

۳-۵-۳-۲-۱ تابع محرک خطی ۴۴
۳-۵-۳-۲-۲ توابع محرک آستانه‌ای دو مقداره حدی ۴۵
۳-۵-۳-۲-۳ تابع محرک زیگموئید ۴۵
۳-۵-۳-۳ مدل چند ورودی ۴۶
۳-۵-۳-۳-۱ فرم خلاصه شده ۴۷
۳-۵-۴ ساختار شبکه‌های عصبی مصنوعی ۴۸
۳-۵-۵ مدل کردن شبکه عصبی مصنوعی ۴۹
۳-۵-۶ کاربردهای شبکه‌های عصبی مصنوعی ۵۰
۳-۵-۶-۱ طبقه بندی، شناسایی و تشخیص ۵۱
۳-۵-۶-۲ پردازش سیگنال ۵۱
۳-۵-۶-۳ پیش‌بینی سری‌های زمانی ۵۲
۳-۵-۶-۴ مدل‌سازی و کنترل ۵۲
۳-۵-۶-۵ بهینه‌سازی ۵۲
۳-۵-۶-۶ سیستم‌های خبره و فازی ۵۲
۳-۵-۶-۷ مسائل مالی، بیمه، امنیت، بازار بورس و وسایل سرگرم کننده ۵۳
۳-۵-۶-۸ ساخت وسایل صنعتی، پزشکی و امور حمل و نقل ۵۳
۳-۵-۷ انواع شبکه‌های عصبی مورد استفاده ۵۳
۳-۵-۷-۱ پرسپترون ۵۳
۳-۵-۷-۲ پرسپترون چندلایه ۵۴
۳-۵-۷-۲-۱ انواع الگوریتم‌های آموزشی مورد استفاده ۵۵
۳-۵-۷-۲-۲ الگوریتم آموزشی لونبرگ مارکوات ۵۵
۳-۵-۷-۳ شبکه‌های توابع پایه شعاعی ۵۶
۳-۶ ماشین بردار پشتیبان ۵۷

۳-۶-۱ مقدمه ۵۷
۳-۶-۲ دلایل استفاده از SVM ۵۷
۳-۶-۳ تعاریف کلی ۵۹
۳-۶-۳-۱ بعد VC ۵۹
۳-۶-۳-۲ حداقل سازی ریسک تجربی ۶۰
۳-۶-۳-۳ حداقل سازی زیسک ساختاری ۶۲
۳-۶-۴ طبقه بندی کننده خطی با داده‌های جداشدنی به طور خطی ۶۳
۳-۶-۵ طبقه بندی کننده خطی با داده‌های جدانشدنی به‌طور خطی ۶۹
۳-۶-۶ طبقه بندی کننده غیر خطی ۷۳
۳-۶-۷ انتخاب تابع کرنل ۷۶
۳-۶-۷-۱ کرنل چندجمله‌ای ۷۶
۳-۶-۷-۲ کرنل شبکه عصبی ۷۶
۳-۶-۷-۳ کرنل گوسی ۷۷
۳-۶-۷-۴ کرنل توابع شعاع مبنا ۷۷
۳-۶-۸ رگرسیون ماشین بردار پشتیبان ۷۷
۳-۷ نرمال کردن داده‌ها ۷۹
۳-۸ داده‌های خروجی ۸۰
۳-۹ تقسیم بندی داده‌ها برای آموزش و آزمون مدل‌ها ۸۰
۳-۹-۱ داده‌های آموزشی ۸۰
۳-۹-۲ داده‌های آزمون ۸۰
۳-۱۰ بررسی اعتبار مدل ۸۱
۴ فصل چهارم: نتایج و بحث ۸۳
۴-۱ مقدمه ۸۴

۴-۲ نتایج تعیین تعداد گام‌های زمانی مؤثر در خروجی ۸۴
۴-۲-۱ نتایج تابع خود همبستگی جزئی سری‌های زمانی آلاینده‌ها با تأخیر (k) ۸۴
۴-۲-۲ نتایج تابع همبستگی عرضی در سری‌های زمانی ورودی با سری زمانی خروجی ۸۶
۴-۳ نتایج مدل‌سازی با شبکه عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان ۸۹
۴-۳-۱ پیش‌بینی بدون استفاده از داده‌های خارجی ۹۰
۴-۳-۲ پیش‌بینی با استفاده سایر داده‌ها ۹۹
۴-۳-۲-۱ پیش‌بینی با استفاده از کل پارامترها ۹۹
۴-۳-۲-۲ پیش‌بنی با استفاده از کل پارامترهای هواشناسی ۱۰۷
۴-۳-۲-۳ نتایج پیش‌بینی با استفاده از کاهش پارامترهای هواشناسی ۱۱۵
۵ فصل پنجم: نتیجه‌گیری ۱۲۴
۵-۱ مقدمه ۱۲۵
۵-۲ نتیجه‌گیری ۱۲۵

فهرست جداول

جدول ۳-۱:موقعیت جغرافیایی ایستگاههای سنجش آلودگی هوا ۳۷
جدول ۳-۲:موقعیت جغرافیایی ایستگاههای هواشناسی ۳۷
جدول ۴-۱: ضرایب همبستگی جزئی سری زمانی آلاینده‌ها تا ۲۴ تأخیر در ایستگاه اقدسیه ۸۵
جدول ۴-۲: ضـرایب همبستگی عرضی سری زمانی پارامترهای ورودی با غلظت SO_2 تا ۲۴ تأخیر ایستگاه اقدسیه ۸۶
جدول ۴-۳: ضـرایب همبسـتگی عرضی سری زمانی پارامترهای ورودی با غلظت NO تا ۲۴ تأخیر ایستگاه اقدسیه ۸۷

جدول ۴-۴: ضرایب همبستگی عرضی سری زمانی پارامترهای ورودی با غلظت NO_2 تا ۲۴ تأخیر ایستگاه اقدسیه ۸۸
جدول ۴-۵: ضـرایب همبستگی عرضی سری زمانی پارامترهای ورودی با غلظت CO تا ۲۴ تأخیر ایستگاه اقدسیه ۸۸
جدول ۴-۶: ضـرایب همبسـتگی عرضی سری زمانی پارامترهای ورودی با غلظت O_3 تا ۲۴ تأخیر ایستگاه اقدسیه ۸۹
جدول ۴-۷: نتایج پیش‌بینی با اسـتفاده از شـبکه‌های عصـبی پرسـپترون، توابع شـعاعی و ماشـین بردار پشتیبان برای ایستگاه اقدسیه ۹۱
جدول ۴-۸: نتایج پیش‌بینی با اسـتفاده از شـبکه‌های عصـبی پرسـپترون، توابع شـعاعی و ماشـین بردار پشتیبان برای ایستگاه پونک ۹۲
جدول ۴-۹: نتایج پیش‌بینی با اسـتفاده از شـبکه‌های عصـبی پرسـپترون، توابع شـعاعی و ماشـین بردار پشتیبان برای ایستگاه ستاد بحران ۹۴
جدول ۴-۱۰: نتایج پیش‌بینی با اسـتفاده از شـبکه‌های عصـبی پرسـپترون، توابع شـعاعی و ماشـین بردار پشتیبان برای ایستگاه شهرداری منطقه ۱۶ ..... ۹۵
جدول ۴-۱۱: تایج پیش‌بینی با اسـتفاده از شـبکه‌های عصـبی پرسـپترون، توابع شـعاعی و ماشـین بردار پشتیبان برای ایستگاه شهر ری ۹۷
جدول ۴-۱۲: نتایج تجزیه واریانس برای غلظت SO_2 با بدون ورودی‌های خارجی ۹۷
جدول ۴-۱۳: نتایج تجزیه واریانس برای غلظت NO با بدون ورودی‌های خارجی ۹۸
جدول ۴-۱۴: نتایج تجزیه واریانس برای غلظت NO_2 با بدون ورودی‌های خارجی ۹۸
جدول ۴-۱۵: نتایج تجزیه واریانس برای غلظت CO با بدون ورودی‌های خارجی ۹۸
جدول ۴-۱۶: نتایج تجزیه واریانس برای غلظت O_3 با بدون ورودی‌های خارجی ۹۸
جدول ۴-۱۷: نتایج پیش‌بینی با اسـتفاده از شـبکه‌های عصـبی پرسـپترون، توابع شـعاعی و ماشـین بردار پشتیبان برای ایستگاه اقدسیه ۱۰۰

جدول ۴-۱۸: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه پونک ۱۰۱
جدول ۴-۱۹: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه ستاد بحران ۱۰۳
جدول ۴-۲۰: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه شهرداری منطقه ۱۶ ..... ۱۰۴
جدول ۴-۲۱: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه شهر ری ۱۰۵
جدول ۴-۲۲: نتایج تجزیه واریانس برای غلظت SO_2 با کل پارامترها ۱۰۶
جدول ۴-۲۳: نتایج تجزیه واریانس برای غلظت NO با کل پارامترها ۱۰۶
جدول ۴-۲۴: نتایج تجزیه واریانس برای غلظت NO_2 با کل پارامترها ۱۰۶
جدول ۴-۲۵: نتایج تجزیه واریانس برای غلظت CO با کل پارامترها ۱۰۷
جدول ۴-۲۶: نتایج تجزیه واریانس برای غلظت O_3 با کل پارامترها ۱۰۷
جدول ۴-۲۷: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه اقدسیه ۱۰۸
جدول ۴-۲۸: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه پونک ۱۰۹
جدول ۴-۲۹: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه ستاد بحران ۱۱۰
جدول ۴-۳۰: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه شهرداری منطقه ۱۶ ..... ۱۱۱
جدول ۴-۳۱: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه شهر ری ۱۱۳

جدول ۴-۳۲: نتایج تجزیه واریانس برای غلظت SO_2 با کل پارامترها هواشناسی..... ۱۱۳
جدول ۴-۳۳: نتایج تجزیه واریانس برای غلظت NO با کل پارامترها هواشناسی ۱۱۴
جدول ۴-۳۴: نتایج تجزیه واریانس برای غلظت NO_2 با کل پارامترها هواشناسی..... ۱۱۴
جدول ۴-۳۵: نتایج تجزیه واریانس برای غلظت CO با کل پارامترها هواشناسی ۱۱۴
جدول ۴-۳۶: نتایج تجزیه واریانس برای غلظت O_3 با کل پارامترها هواشناسی..... ۱۱۴
جدول ۴-۳۷: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه اقدسیه..... ۱۱۶
جدول ۴-۳۸: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه پونک..... ۱۱۷
جدول ۴-۳۹: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه ستاد بحران..... ۱۱۸
جدول ۴-۴۰: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه شهرداری منطقه ۱۶..... ۱۱۹
جدول ۴-۴۱: نتایج پیش‌بینی با استفاده از شبکه‌های عصبی پرسپترون، توابع شعاعی و ماشین بردار پشتیبان برای ایستگاه شهر ری ۱۲۰
جدول ۴-۴۲: نتایج تجزیه واریانس برای غلظت SO_2 با کل کاهش پارامترها هواشناسی..... ۱۲۱
جدول ۴-۴۳: نتایج تجزیه واریانس برای غلظت NO با کل کاهش پارامترها هواشناسی..... ۱۲۱
جدول ۴-۴۴: نتایج تجزیه واریانس برای غلظت NO_2 با کل کاهش پارامترها هواشناسی..... ۱۲۱
جدول ۴-۴۵: نتایج تجزیه واریانس برای غلظت CO با کل کاهش پارامترها هواشناسی..... ۱۲۲
جدول ۴-۴۶: نتایج تجزیه واریانس برای غلظت O_3 با کل کاهش پارامترها هواشناسی ۱۲۲
جدول ۴-۴۷: درصد فراوانی مدل‌های پیش‌بینی کننده در ورودی‌های مختلف ۱۲۳

فهرست شکل‌ها

شکل۱-۱: ساختار کلی شبکه عصبی مصنوعی ۱۶

شکل ۳-۱: سیستم عصبی ۴۰

شکل ۳-۲: مدل نورون تک ورودی ۴۳

شکل ۳-۳ تابع محرک خطی ۴۴

شکل ۳-۴: تابع محرکه آستانه‌ای دو مقداره حدی ۴۵

شکل ۳-۵: تابع محرک زیگموئید ۴۶

شکل ۳-۶: شکل چند ورودی چند نورون ۴۷

شکل ۳-۷: فرم خلاصه شده نورون با R ورودی ۴۸

شکل ۳-۸: شبکه عصبی توابع شعاعی ۵۶

شکل ۳-۹ سه نقطه در فضای ئو بعدی که خطوط جهت دار از هم نفکیک شده‌اند ۶۰

شکل ۳-۱۰: فوق صفحه جدا کننده در حالت بطور خطی جداپذیر. دایرهها بردار پشتیبان ۶۵

شکل ۳-۱۱ نمونه صفحه جداکننده بهینه (سمت راست). صفحه جداکننده دو کلاسه (سمت چپ) ۶۶

شکل ۳-۱۲: فوق صفحه جدا کننده برای حالت جدانشدنی به طور غیر خطی ۶۹

شکل ۳-۱۳: نگاشت داده‌های آموزشی غیر خطی به فضایی از ویژگی‌ها با ابعاد ۷۴

فصل اول

مقدمه و کلیات

۱-۱ مقدمه

موضوع کیفیت هوا از موضوعات اصلی موردبحث و تحقیق در میان سیاستگزاران و مردم عامه است. آلودگی هوا یک پدیده هواشناسی ناشی از فعالیت‌های انسان برای رسیدن به حداکثر منافع خود و بدون توجه به‌سلامت طبیعت است و برای بسیاری از افراد به یک موضوع نگران‌کننده مبدل شده چراکه به‌طور مستقیم سلامت انسان را به خطر می‌اندازد (به‌عنوان‌مثال مشکلات تنفسی، بیماری‌های قلبی و ریوی را به همراه دارد و حتی ممکن است سبب مرگ زودرس شود). کودکان به نسبت فعالیت بیشتری در هوای آزاد و خارج از منزل داشته و ریه‌های آن‌ها هنوز در حال رشد و توسعه هستند، بنابراین بیشتر درمعرض خطرند، ضمن اینکه افراد مسن نیز به برخی از انواع آلودگی هوا حساس می‌باشند. تغییر در میزان کیفیت هوا از منابع مختلف آلاینده که عمدتاً شامل گازهای صنعتی ناشی از کارخانه‌ها و کارگاه‌ها، منابع متحرک از قبیل اتومبیل، اتوبوس، کامیون، هواپیما و کشتی؛ و همچنین منابع زیستی مانند آتش‌سوزی و گیاهان و گردوغبار معلق در باد هستند ناشی می‌شود. میزان انتشار آلاینده‌ها در اتمسفر و حذف آن‌ها می‌تواند تحت تأثیر عواملی چون شدت انتشار از منبع آلودگی، نور خورشید، جغرافیا، رطوبت، ابرها، بارندگی و الگوهای آب و هوایی محلی و منطقه‌ای قرار گیرد. درواقع حوادث آلودگی هوا غالباً در شرایط جغرافیایی و هواشناسی که اجازه گردش سلیس هوا را نمی‌دهد و بخش عظیمی از جمعیت بین مکان‌های دور یک شهر جابجا می‌شوند رخ می‌دهد. این وقایع نیاز به اقدامات شدید از قبیل بسته شدن مدارس و کارخانه‌ها و محدود نمودن تردد خودروها دارد لذا پیش‌بینی آلودگی اجازه در نظر گرفتن اقدامات کارآمدتر برای حفاظت از سلامت شهروندان را می‌دهد. وضع قوانین در خصوص آلودگی هوای شهری و برون‌شهری نیاز به آنالیز و پیاده‌سازی روش‌های عامل خودکار به‌منظور جلوگیری از به آستانه خطر رسیدن آلاینده‌های اصلی هوا دارد. برای رسیدن به این مقصود نیاز ب صرف هزینه و حمایت از تحقیقات پیش‌بینی آلودگی هوا و فراهم نمودن رویکردی استاندارد و معیار در تبادل

داده‌ها و قابلیت‌های چند مدلی برای پیش‌بینی کیفیت هوا است.

۱-۲ بیان مسئله

هوا به‌عنوان یک تعریف مصطلح برای توصیف مخلوطی از گازها که هر یک قشر نسبتاً نازکی را در اطراف زمین به وجود می‌آورند به‌کاربرده می‌شود. ترکیب این مواد، از سطح زمین به‌طرف بالا تا حدود ۱۰ کیلومتر، به‌طور قابل‌ملاحظه‌ای ثابت است. بنابراین آلودگی هوا عبارت است از هرگونه تغییر در ویژگی‌های اجزاء متشکل هوا به‌طوری‌که استفاده پیشین از آن ناممکن گردد و به‌طور مستقیم و غیرمستقیم منافع و حیات موجودات زنده را به مخاطره اندازد. ذرات معلق، به‌عنوان یکی از شاخص‌های آلایندگی هوا، نقش مهمی در آلودگی هوای شهری دارند و از منابع طبیعی شامل طوفان گردوغبار، فرسایش خاک، آتش‌سوزی جنگل‌ها، آتش‌فشان‌ها و همچنین منابع مصنوعی ازجمله احتراق سوخت، حمل‌ونقل و فرآیندهای صنعتی انتشار می‌یابند[1]. امروزه مشکلات ناشی از آلودگی هوا در کلان‌شهرها به یک چالش زیست‌محیطی اساسی تبدیل‌شده است. این مشکل در مورد شهر تهران به دلیل حجم ترافیکی سنگین ناشی از ترابری، استفاده از خودروهای غیراستاندارد، احتراق ناقص سوخت‌های مورداستفاده خودروها و بی‌توجهی که طی سال‌های گذشته نسبت به آلودگی هوا در آن صورت گرفته، اهمیت ویژه‌ای دارد[2]. آلودگی محیط که منظور آلودگی هوا، آب‌وخاک است، امروزه مشکل بزرگی را به وجود آورده است[3]. آلودگی هوا به سبب نیاز مبرم تمام موجودات و ازآن‌جهت که مرز نمی‌شناسد و دیگر آلودگی‌ها ازجمله آلودگی خاک و آب را سبب می‌گردد دارای اهمیت فروانی می‌باشد[4]. هوا و آثار زیان‌بار و فاجعه‌آفرین آن بر کسی پوشیده نیست[5]. به‌طوری‌که هرسال علت مرگ حدود۳-۲/۷ میلیون نفر (۶ درصد از مرگ‌ومیرهای سالیانه) به آلودگی هوا نسبت داده می‌شود[6] و هزینه صرف شده سالیانه برای بخش سلامت و بهداشت ناشی از آلودگی هوا در اتریش، فرانسه و سوئیس به‌عنوان نمونه حدود ۳۰ میلیارد پوند بوده است[7]. در اغلب شهرهای بزرگ و صنعتی دنیا، افزایش جمعیت و منابع آلاینده هوا منجر به تولید هوایی آلوده و ناسالم شده است که سلامتی افراد ساکن در این مناطق را در معرض خطر قرار می‌دهد. در کلان‌شهر تهران نیز افزایش زندگی ماشینی و فعالیت‌های افراطی

انسان ازیک‌طرف و شرایط توپوگرافی و عوامل طبیعی آن از طرف دیگر باعث شده است تا این شهر از آلوده‌ترین شهرهای جهان محسوب شود . به دلیل ضرورت جلوگیری و کاهش خطرهای ناشی از آلودگی هوا، آگاهی مناسب از جوانب مختلف این مسئله از اهمیت بسزایی برخوردار است[8].

مدل، واژه‌ای با کاربرد تقریباً روزمره برای تمامی افراد است. نگاهی واژه شناسانه به جنبه‌های معنایی مدل، حاکی از وجود بیش از هشت معنای متفاوت برای مدل و مدل‌سازی است. صرف‌نظر از معانی مختلف، جوهره اصلی تمامی تعاریف ارائه‌شده برای واژه مدل، تلخیص، ساده‌سازی و کاهش پیچیدگی‌ها است. به‌طورمعمول، یک مدل عبارت از واقعیت ساده‌شده‌ای است که درک و شناخت آن تسهیل بخشیده شده باشد[9]. امروزه پیش‌بینی آلودگی هوا در نواحی شهری به دلیل تأثیر آن بر روی سلامتی انسان، یکی از موضوعات مهم در تحقیقات زیست‌محیطی می‌باشد. غلظت‌های با آلودگی بالا، تأثیرات سوء و مرگ زودرس گروه‌های حساس و آسیب‌پذیر جامعه ازجمله افراد مسن و کسانی که تنگی نفس دچار هستند را به دنبال دارد[10]. پیش‌بینی غلظت روزانه آلاینده‌های هوا اولین گام اساسی در برنامه‌ریزی کاهش اثرات آن‌ها است. بدین منظور تاکنون روش‌های زیادی برای پیش‌بینی غلظت آلاینده‌های هوا مطرح‌شده است که آن‌ها را می‌توان به دودسته قطعی و آماری تقسیم کرد. توسعه مدل‌های قطعی برای مدل‌سازی کیفی آلودگی هوا پیچیدگی‌های زیادی دارد و ایجاد روابط بین تولید و انتشار آلاینده کار چندان ساده‌ای نیست. این عامل باعث وابستگی زیاد این مدل‌ها به تحلیل دقیق انتشار و نیز هواشناسی شده است که به‌نوبه خود عدم قطعیت قابل‌توجهی دارند. روش‌های آماری با استفاده از داده‌های موجود هواشناسی و آلودگی و تحلیل ارتباط آماری بین آن‌ها، راهکارهای ساده‌تری برای پیش‌بینی غلظت آلاینده‌ها به شمار می‌روند و تحقیقات صورت گرفته نیز درزمینه پیش‌بینی کوتاه‌مدت آلاینده‌های هوا با استفاده از روش‌های آماری، سودمندی این روش‌ها را به اثبات رسانده است[11]. روش‌های آماری علاوه بر اینکه به اطلاعات انتشار و ضرایب انتشار، که دسترسی به آن‌ها در بیشتر موارد با مشکلاتی همراه است، نیازی ندارند، نسبت به مدل‌های قطعی ساختار ساده‌تری نیز دارند. روش‌های آماری متعددی برای پیش‌بینی غلظت آلاینده‌های هوا وجود دارد. از میان این روش‌ها تابه‌حال مدل‌های رگرسیون